Invited Paper

Measuring the Size of Lepidopteran Populations

Lawrence F. Gall

Department of Biology and Peabody Museum of Natural History, Yale University, New Haven, Conneticut 06511

Introduction

When out in the field, do you often ask: "how many individuals of this butterfly or moth species are present here?" This number, the population size, is of great interest to the casual observer, collector, and research biologist alike (for different reasons)—but the simplicity of your question is deceptive. More often than not, a satisfactory numerical answer is difficult to obtain. The field of mark, release, and recapture (MRR) has developed to provide answers to such questions about animal abundance.

This paper is a largely non-technical synopsis of MRR theory and practice as usually applied to Lepidoptera. I will first describe several common MRR models, including the assumptions and the formulae with which each calculates population size; I then cover how they are put to practice, some associated pitfalls in interpretation, and the relative merits of the different models. Mark-recapture models estimate *absolute* abundance i.e., the actual number of individuals present. Because (1) absolute estimates may not be one's primary interest, and (2) conducting a mark-recapture study on a butterfly or moth population can be a time consuming venture, or even impossible, I will also discuss some simpler non-marking techniques for estimating *relative* abundance.

For the novice and sophisticated reader alike, I highly recommend Begon's (1979) little paperback, *Investigating Animal Abundance*, as a precis on mark-recapture. He presents the principles lucidly, and offers some of the best available advice for data analysis and interpretation. Blower et al.'s (1981) book is another good introductory reference. For the detailed statistical properties of MRR models, consult the treatises by Seber (1973) and especially Cormack (1968, 1979).

ABSOLUTE ABUNDANCE MODELS

Marking

To use most MRR models one must assign at minimum a *date-specific* mark to each captured animal. The main exception is the Lincoln Index

(and its derivatives), in which there is a single marking period, and animals thus must only be classifiable upon later recapture as either marked or unmarked. Nevertheless, it is always best to assign a *unique mark* to each individual, because other factors (e.g., flight distances and speeds) are hard to quantify unless one can recognize individual animals.

How does one mark butterflies and moths? In virtually all cases, the simplest mark to apply will be a number. Even with species having small wingspans, a unique number can be written directly on one or more wings (usually the basal underside) using fast-drying, permanent ink; "Sharpie" felt-tip pens or their equivalent serve admirably. For species with variegated wing patterns, Ehrlich and Davidson's (1960) "1-2-4-7" system or its modification (Brussard, 1970) can be used; Southwood (1978) summarizes the diversity of other available marking methods and coding schemes. For lepidopteran work in general, I advise against codes, for two reasons: (1) numbers are simpler to write and remember; (2) codes are very easily misread. Other aspects of the marking process are treated elsewhere in this paper.

Principles

All absolute abundance MRR models share a common array of assumptions, and most models also make additional ones. These assumptions are interrelated, and encompass the many subtle aspects of physical and temporal patterns of sampling, and the behavior of the animals themselves. The major ones are:

1) sampling is done in discrete intervals, that are short in relation to the total time of the study
2) marked animals do not lose their marks
3) marked animals can be distinguished from unmarked ones
4) once marked, the behavior (*sensu lato*) of animals does not change
5) marked animals mix thoroughly with the unmarked animals
6) marked and unmarked animals have the same probability of capture
7) sampling is random with respect to mark status (i.e., sexes, age classes, etc. are sampled at their natural proportions)

How to deal with these assumptions is covered in greater depth later.

I discuss four absolute abundance MRR models in this section: the Lincoln Index, Fisher-Ford, Jolly-Seber, and Manly-Parr models. These and other models can be categorized broadly as either "single-marking" or "multiple-marking" models. Single-marking models consist of one sample during which marking and release are conducted, followed by one or more samples in which animals are recaptured (the Lincoln Index is a single-marking model). Multiple-marking models consist of a series of sampling

periods, during each of which marking, release, and recapture are conducted (Fisher-Ford, Manly-Parr, and Jolly-Seber are multiple-marking models).

The Lincoln Index, Fisher-Ford, Jolly-Seber, and Manly-Parr models determine population size using *ratios of marked to unmarked individuals.* The basic sampling principle is easily illustrated. Imagine a large box filled with ping-pong balls. A sample of balls is taken from the box, and each is marked with a stripe. The striped balls are returned to the box, which is then shaken vigorously. A second sample of balls is now drawn, which (in all likelihood) will contain some striped balls. The estimate of the total number of balls in the box then is: the number of balls in the first sample divided by the proportion of striped balls in the second sample (exactly so for the Lincoln Index—the three other models use only subtly different ratios to determine the total).

Single-Marking Models: the Lincoln Index

The Lincoln Index, or Peterson Estimator, is probably the most familiar of all absolute abundance models, and is the simplest and oldest of those described in this paper. Only two samples are necessary to obtain a Lincoln Index: a marking and recapture sample. The ping-pong example above calculated a simple Lincoln Index. In general, with:

n_1 = number of animals marked and released in first sample
n_2 = number of animals captured in the second sample
m = number of marked animals in the second sample
$\hat{N}$ = total population size

$$\hat{N} = \frac{n_1 \, n_2}{m} \quad \text{with} \quad \text{VAR } \hat{N} = \frac{n_1 \, n_2 \, (n_2\text{-}m)}{m}$$

Bailey (1951, 1952) showed that this form of the Lincoln Index has a positive bias of the order $1/m$, and proposed the following continuity correction:

$$\hat{N} = \frac{n_1 \, (n_2+1)}{m+1} \quad \text{with} \quad \text{VAR } \hat{N} = \frac{n_1 \, (n_2+1) \, (n_2\text{-}m)}{(m+1) \, (m+2)}$$

Estimates of population size using either this or the first formulation differ only slightly unless sample sizes are very small.

Peterson, in 1889, and Lincoln, in 1930, are usually cited as the first to use the method. Bailey (1952) and LeCren (1965) both noted that Peterson apparently used his fish recaptures just to calculate mortality rates, and Lincoln's waterfowl study thus has priority in actual application of the model logic for estimating population size. However, LeCren (1965) also pointed out that Dahl used the principle in his 1917 studies of trout, and Cormack (1968) credits Sir Francis Bacon with similar reasoning several centuries earlier still. Because the method is so intuitive, there is

no doubt it has been independently discovered many other times.

The advantage of the Lincoln Index is its ease of calculation. Its disadvantage is an additional assumption, which rarely if ever will hold for the majority of lepidopteran populations:

8) the population suffers no gains or losses during the sampling interval.

Gains include birth (recruitment) and immigration; losses include mortality and emigration. Assumption eight thus addresses the concept of *population closure*. Actually, a population need not be truly closed to use the Lincoln Index: the model can be applied if there is neither recruitment nor immigration (which affect the number of unmarked animals), and if mortality and emigration also affect the marked and unmarked animals equally. Alternatively, one can make independent estimates of gains and losses, and account for them during data analysis (appropriately modified formulations of the Lincoln Index are discussed by Seber, 1973; and Begon, 1979, describes a weighted-mean version of the Lincoln Index for use with several days' recapture data). The Schnabel census and Jackson's "positive method" (see Cormack, 1979) are also essentially Lincoln Index models, the former being multiple-marking, the latter single-marking.

Multiple-Marking Models

The Fisher-Ford, Manly-Parr, and Jolly-Seber models all offer significant improvement over the Lincoln Index, by accounting for some or all of the sources of population gains and losses. Assumption eight is thereby relaxed. These three models require a series of censuses to determine gains and losses. Because previously marked animals are recaptured (and possibly remarked) on the second and later samples, multiple-marking models assume that:

9) the probability of recapture is unaffected by the number of previous captures (simply an extension of assumptions four and six).

FISHER-FORD

The Fisher-Ford model (Dowdeswell et al., 1940), Schnabel census, and Jackson's "negative method" are all early contemporary multiple-marking models. Fisher-Ford is sometimes referred to as a "trellis" model, for the manner in which the raw data are set up to do the calculations. Using terms as before, and with:

$\emptyset$ = residence rate over the period i to t (probability that an animal present at time i will be present at time t)

r = recaptures at time t of animals marked at time i

and: $\hat{N} = \dfrac{n_t \, m_i \, \emptyset}{r}$

(The symbol Ø is often referred to as the survival rate, but the term residence is more appropriate, since in practice death and emigration can usually not be separated from one another as sources of population loss). A *constant* residence rate which empirically best fits the data is first found by trial and error (procedure outlined by Fisher and Ford, 1947; and Begon, 1979). This rate is then used to determine the necessary Ø s.

A particular advantage of the Fisher-Ford model is that other variables e.g., catch rate, or periodicities in animal activity, can be incorporated easily into the model structure (Seber, 1973; Southwood, 1978). The (debatable) disadvantage is the assumption of a constant residence rate throughout the course of the study (see below). Bailey's (1952) Triple-Catch is a special three-sample case of the Fisher-Ford model.

JOLLY-SEBER

The Fisher-Ford model and its contemporaries are *deterministic*—one or more parameters are invariant. For example, an individual's chance of surviving from one sample to the next is assumed in Fisher-Ford to be an exact value, rather than a probability. Both the Jolly-Seber and Manly-Parr models are *stochastic*—the model parameters represent probabilities. Jolly (1965) and Seber (1965) independently derived stochastic models for open populations, although Jolly's differs by allowing for removal of captured animals from the population, an important consideration with Lepidoptera. In the terms of Jolly (1965), let:

n_i = number of animals captured in the ith sample
m_i = number of previously marked animals in the ith sample
s_i = number released from the ith sample after marking
r_i = number of the s_i which are caught subsequently
z_i = number of animals marked before time i which are *not* caught in the ith sample, but which *are* caught subsequently

$$\hat{N}_i = \frac{\hat{M}_i\, n_i}{m_i} \quad \hat{M}_i = m_i + \frac{s_i z_i}{r_i} : \quad \hat{N}_i = n_i + \frac{n_i s_i z_i}{m_i r_i}$$

$\hat{M}_i$ is an estimate of the total number of marked animals "at risk" of capture in the population at time i. The variance formula is:

$$\text{VAR}_{\hat{N}_i} = \hat{N}_i(\hat{N}_i - n_i) \left[\frac{\hat{M}_i - m_i + s_i}{\hat{M}_i} \left(\frac{1}{r_i} + \frac{1}{s_i} \right) + \left(\frac{1}{m_i} - \frac{1}{n_i} \right) \right]$$

The advantage of the Jolly-Seber model over the previous models is that it is fully stochastic, and can account for the usual sources of population gains and losses. It requires at least three samples, and makes an added assumption:

10) the probability of surviving from one sample to the next is the same

for each marked animal (i.e., age-independent residence)

The Jolly-Seber model has been criticized because it requires many parameters to be estimated, and is thus not parsimonious (Cormack, 1979). Various authors have also expressed the need for a stochastic model which assumes a (biologically often justifiable) constant residence rate, as with Fisher-Ford. Accordingly, Jolly (1982) recently developed three modified versions of the original Jolly-Seber model, covering the situations where residence rate, probability of capture, and both parameters are constant over time. Following Seber (1973), Jolly (1982) also adopted the following continuity correction for the original 1965 model, as per Bailey's (1952) modification of the Lincoln Index:

$$\hat{N}_i = \frac{\hat{M}_i \, n_i}{m_i} \qquad \hat{M}_i = m_i + \frac{(s_i + 1) \, z_i}{(r_i + 1)} : \qquad \hat{N}_i = n_i + \frac{(n_i + 1)(s_i + 1) \, z_i}{(m_i + 1)(r_i + 1)}$$

MANLY-PARR

Manly and Parr (1968) noted that with short-lived, discrete generation insects the age of animals marked first on day i is likely to be less than that of animals marked before day i. Because of the concomitant possibility of age-dependent mortality, and hence violation of assumption ten above, Manly and Parr developed a model based on sampling intensity—$n_i/\hat{N}_i$—a measure of the fraction of the resident population processed on a sampling occasion. The best available estimate of sampling intensity is $m_i/\hat{M}_i$, which can be calculated by setting up an individual mark table. For each day i, assign to an animal one of the symbols:

x = if this is its first or last capture
y = if this is an intermediate capture
z = if it is *not* captured, but *is* known to be present

Then, for any day i, the estimated sampling intensity will be the number of marked animals captured on day i divided by the number known to be present before and after, thus:

$$\hat{N}_i = \frac{n_i}{p_i} \qquad \text{where sampling intensity, } p_i = \frac{m_i}{\hat{M}_i} = \frac{\Sigma \, y_i}{\Sigma \, y_i + \Sigma \, z_i}$$

$$\text{or:} \quad \hat{N}_i = \frac{n_i \, (\Sigma \, y_i + \Sigma z_i)}{\Sigma \, y_i}$$

The Manly-Parr model relaxes the assumption about mortality being independent of age. Its disadvantage is that it requires a high sampling intensity; Manly and Parr felt that y_i should exceed 10 for the method to be considered reliable.

Other Absolute Abundance Models

FREQUENCY-OF-CAPTURE

Some absolute abundance MRR models are not based on ratios of marked to unmarked individuals. Perhaps the best known are the "frequency-of-capture" models, which rely upon the distribution of different recapture classes: i.e., the number of animals caught once; number caught twice; thrice; and so forth. The "zero class" is the number of animals never caught, and the sum of all classes is the total population size. Various truncated discrete probability distributions have been fitted to the observed distributions of recapture classes, including the binomial, Poisson, and geometric distributions (see Seber, 1973, and Caughley, 1977).

Craig (1953) developed a frequency-of-capture model specifically with butterflies in mind. With the following terms:

x = number of times an individual has been marked
f = number of individuals that have been caught x times
$\sum xf$= total number of capture events (1 times number caught once, plus 2 times number caught twice, and so forth)

$$\text{and:} \quad \hat{N} = \frac{(\sum xf)^2}{\sum x^2 f - \sum xf}$$

Another method employing the same terms is that of Edwards and Eberhardt (1967), who used their model to measure cottontail abundance:

$$\hat{N} = \frac{\sum f}{1 - (\sum f / \sum xf)}$$

Frequency-of-capture models are often applied in a single-marking fashion, with one marking census and several recapture censuses (as per Edwards and Eberhardt, 1967). With highly mobile animals, these census periods can in principle be collapsed. The attraction that frequency-of-capture models thus have for work with Lepidoptera is that an absolute population estimate can be provided for a very brief time period—a day, or even less (Craig, 1953, envisioned a one day census). Problems with compressing the sampling interval are pursued to some extent later; essentially, immediately upon release, one must presume that marked animals are (1) catchable again, and (2) have mixed back into the population. The analogy is marking single ping-pong balls from the box described earlier, being able to shake the box vigorously enough in a second or so to mix all the balls, and repeating the sampling procedure many times for, say, an hour.

REMOVAL

Seber (1973), Caughley (1977) and Southwood (1978) all review models which rely upon removing segments of the population to estimate population size. The simplest technique is, on each of several sampling occasions, to capture a series of animals and *not* release them back into the population. The rate at which successive sample sizes drop off is proportional to the total population size and the number removed. The number of animals removed on the ith sample can be plotted (y-axis) against the sum of animals removed before the ith sample (x-axis), and the total population size (x-intercept) determined by linear regression.

A variant of this approach is to employ the ratios of "natural marks" in a population e.g., males and females, or polymorphs. The proportions of the natural marks are determined in a prior survey, and then removal sampling on one of the mark classes is carried out. The change in ratio of the natural marks from the first sample to the second is related to the total population size. This is Kelker's (1940) "change-in-ratio" method:

x = mark class from which individuals are removed
y = mark class not removed
n = *number* of the x that are removed
p = *proportion* of mark class i in population at time j

$$\widehat{N} = n \left[p_{x1} - \frac{p_{y1}\, p_{x2}}{p_{y2}} \right]^{-1}$$

Dealing with the Assumptions

Failure to meet one or more assumptions made by a MRR model leads to predictably negative results. If one is unaware of violations, then there will more than likely be serious errors in interpretation. If one knows of the violations, then how one *tempers* data interpretation is the primary concern (it should go without saying that if assumptions are violated, one must not interpret results as if they held). One can often make appropriate allowances for violated assumptions when calculating populational parameters.

It is worth briefly listing the effects that violating three assumptions have on estimates of population size. Complete discussion of all aspects of MRR assumptions are given by Cormack (1979) and Begon (1979; Chapters 3-4). The three assumptions treated here are composites of the ten enumerated earlier; in practice, these three will also likely be confounded.

MARKING DOES NOT ALTER SUBSEQUENT ACTIVITY

A principal assumption of all but removal models is that marking does

not alter patterns of activity. Marking may either diminish an individual's chance of subsequent capture ("trap-shyness") or enhance its likelihood ("trap-happiness"); and with social animals, marking may even influence the unmarked members of the population.

If the marking process decreases the probability of recapture, then the number of marked animals in subsequent samples will be underrepresented, and $\hat{N}_i$ accordingly overestimated. If marked animals are more likely to be captured later, the converse is true. In *open* populations, $\hat{N}_i$ is unaffected if marking alters the probability of survival (perhaps a counterintuitive result). In *closed* populations, a decrease in survival probability leads to an overestimate in $\hat{N}_i$; conversely with an increase in survival probability.

One of the few times a removal model has an advantage over other MRR models is when marking *does* influence activity (the other time is as an alternative to frequency-of-capture models during a single census). Since there are no releases using removal methods, one is free from the assumption that marking has no effect. Removal sampling of Lepidoptera can be carried out non-destructively by accumulating all captured individuals in a flight cage until all sampling is complete, at which point releases are made. However, except during quite restricted instances, removal models will be inferior to the other MRR models. I pursue the assumption that marking does not alter behavior again in the context of comparing absolute and relative abundance models.

THE POPULATION IS CLOSED

Closure implies that there are neither gains (births, immigrations) nor losses (deaths, emigrations). The Lincoln Index and frequency-of-capture models assume population closure. When both gains and losses occur, the number of marked animals is being diluted over time. Overestimates of $\hat{N}_i$ will thereby occur when using these models. The same result can be expected when there are only gains to the population. If, however, there are only losses, *and* these occur in similar frequency in both the marked and unmarked fractions of the population, then $\hat{N}_i$ remains unaffected.

Implicit in the concept of closure is that emigration is permanent. If an animal emigrates from a population but returns again much later, then in practice it has been "trap-shy," as with an animal whose probability of recapture was decreased by marking. The effects on $\hat{N}_i$ estimates are then the same as those described in the preceding section.

ALL INDIVIDUALS ARE EQUALLY CATCHABLE

Most models assume that all animals in a population have the same catchability. There are many instances where catchabilities might differ: for example, inactive individuals are less likely to be captured than active

ones; and dominant individuals may be more visible than subordinate ones.

Should certain individuals be consistently more catchable than others, then the population size will be underestimated. If, however, the catchability differences are not consistent from census to census, the $\hat{N}_i$s will remain largely unaffected. A common situation is systematic differences in catchability among the sexes, or age classes; population size will tend to be underestimated in such instances.

"Probability of recapture" and "catchability" are potentially confusable terms. Strictly, at time i, the probability of recapture is the product of catchability (p_i) times residence ($\emptyset_i$). Because these two biologically distinct factors can mask one another, differences in recapture probability must be interpreted with caution. Working with *Colias* butterflies, Tabashnik (1980) developed new methods for disentangling these two elements of recapture probability; I have also explored the usefulness of these tests using data from *Boloria* (Gall, 1984a, 1984b). Because Tabashnik's methods for partitioning the components of recapture probability offer improvement over earlier ones (see Begon, 1979), I feel that his two tests should be incorporated as a matter of course into mark-recapture studies. Interested readers should consult Tabashnik (1980) for details; Carothers (1973) is another most illuminating paper that treats the catchability of taxi-cabs.

Utility of the Different Models

I have noted one or several circumstances in which each MRR model may be considered particularly appropriate. How do these models perform when pitted against each other? When making such comparisons, it is important to bear in mind that each model is designed for particular circumstances (hence the variation in assumptions). The performance of any model therefore is strongly study-dependent; different organisms and situations may define mutually exclusive sets of models as choices. For example, when a population suffers losses and gains, one would not select the simple Lincoln Index or a frequency-of-capture model, because these depend upon closure. Such models are designed for other situations. Whenever possible, though, one should compare results from an array of different MRR models, even if one or more of the models appears less appropriate *a priori*, because comparison provides crucial insight into the processes operant in the population, and helps clarify model applicability.

The Fisher-Ford, Manly-Parr, and Jolly-Seber models are considered to have the broadest applicability among available MRR models. The 1965 Jolly-Seber model is also currently touted as the brand leader among them, evidence having accumulated now from simulation studies, field work, and combined approaches (e.g., Manly, 1970; Bishop and Shep-

pard, 1973; Roff, 1973; Cormack, 1979; Begon, 1979; and references therein). Jolly's (1982) revised model, allowing for constant residence rate and/or probability of capture, will no doubt prove more broadly applicable still (few studies using the new formulations have yet been reported). These are vigorous endorsements for the Jolly model, but the salient points to remember in comparison are: Fisher-Ford assumes the most but requires the least data; Manly-Parr assumes the least but requires the highest sampling intesities; Jolly-Seber is intermediate on both counts. Begon (1979, pp. 53-54) summarizes:

> "If the data are sparse, and survival-rate both constant and age-independent, then Fisher-Ford is obviously the most applicable method. If the data are extensive, and survival-rate both variable and age-dependent, then Manly-Parr is appropriate. But there will be many situations in which the pros and cons are shared more evenly. It should be noted, for instance, that the more restrictive models are both fairly robust when their assumptions are violated. Thus, Jolly is preferable to Fisher-Ford only if survival-rate varies *significantly*, and Manly-Parr preferable to Jolly only if survival is *strongly* age-dependent."

I stress the distinction between concluding that a model is more applicable, and concluding that it is in some intrinsic sense "better" than others. Because model performance depends upon context, the latter conclusion does not necessarily follow.

RELATIVE ABUNDANCE MODELS

There are at least three cogent reasons why one might opt *not* to conduct a mark-recapture program to assess the abundance of a particular species. First, the cost of gathering such data is high. One typically must invest a large amount of energy in both field work and analysis time to carry out an MRR stiudy. Second, it may be impossible to conduct an MRR study: individuals may fly at the tops of trees; netting specimens may not be allowed; and so forth. Third, *absolute* abundance may not be of primary interest. The investigator may be asking: how has the size of the butterfly population in my backyard varied over the past ten years? Such a question deals with *relative trends* in abundance, and does not strictly require that absolute numbers be known.

Overview

Most techniques for measuring relative abundance are simple, requiring a minimum of investigator effort and equipment. Arrays of relative methods are documented in the literature, many having been conceived for a single species (and/or out of necessity, because applying an absolute abundance model would have been out of the question). These span the range from listening for animals, through direct counts and transect

sampling, to both passive and mobile traps; and also include methods based not on the animals themselves, but on by-products such as feces or extent of defoliation (see Doane and McManus, 1981, for examples of the use of by-products as indices of lepidopteran populations). Southwood (1978) provides a concise and sometimes amusing synopsis of relative abundance methods with particular reference to insects.

The measurement of relative abundance is enjoying a recent surge in popularity. This stems directly from increased focus on global conservation issues, and the corollary gathering of long-term data on populations. Discussion of relative abundance methods as applied to Lepidoptera is appropriately set in such a conservational context.

Butterfly Counts, and Lepidoptera Conservation

Britain has a long-standing commitment to conserving Lepidoptera and their habitats, and their techniques for studying changes in butterfly and moth populations are accordingly well-developed. Researchers in Britain recognized the "need for a simple reliable method of recording abundance of butterflies in nature reserves and similar places so that changes from year to year can be assessed" (Pollard et al., 1973, p. 79; see also the Scandinavian work by Douwes, 1970, 1976). Out of this need grew the butterfly count—a relative abundance method, broadly defined as a census in which one records the number of individuals seen of different species, according to some predefined spatial and/or temporal rules.

The British have gradually settled on line-transects for butterfly counts, the most prevalent one being that used by the Institute for Terrestrial Ecology (ITE). Briefly, the recorder walks along a pre-determined linear path at a uniform pace, counting butterflies within 5 meters. Counts are made between 1045 and 1545 hours BST; counts are not made when the temperature is below 13°C, only in sunny conditions between 13-17°C, and in any condition other than inclement weather above 17°C. The mean count per transect can be determined on a weekly basis, and these weekly means summed over the entire brood to give an index of abundance. The ITE procedure is an extension of the count technique used in the 1960s by Moore, and is described in depth by Pollard (1977). Thomas (1983) recently established a routine for standardizing counts to allow site-to-site comparisons; let:

L = length of the count transect, in meters
A = size of flight area, in hectares
N = butterfly numbers per 100 meters of transect

$$\text{population index, } P = \frac{100\ N\ A}{L}$$

Ambitious monitoring programs overseen by ITE using the line-transect are underway in Britain; summaries of count results, and their use in

habitat management and ecological research, have been published widely (Pollard, 1977, 1979, 1984).

In North America in 1975, the Xerces Society established its Fourth of July Butterfly Count (4JBC), based on the success of both the British experience and the Christmas bird count of the Audubon Society (Hughes, 1975; the ITE count ultimately also harkens to the Common Bird Census of the British Trust for Ornithology). The 4JBC is a single-day count of butterflies in a circular area, 7.5 miles in diameter from an established central point. The count procedure is fundamentally similar to that described for the ITE line-transect (see Opler and Powell, 1984).

In contrast to the ITE counts, little has yet been done formally with the 4JBC database. Because the 4JBC is only a one day count, the data will obviously be less sensitive to trends than if the censuses were more frequent. But daily/weekly butterfly counts are established in several areas, and studies of trends in relative butterfly abundance in North America are thus becoming available (e.g., Smith, 1984). In general, lepidopteran conservation in the Nearctic is still gathering steam (reviews by Pyle, 1976, and Pyle et al., 1981). Notably, only recently have extensive mark-recapture studies aimed specifically at conserving Nearctic Lepidoptera reached the primary ecological literature (see Arnold, 1983, for lycaenids and *Speyeria* in coastal California; Gall, 1984a, 1984b, for the endemic *Boloria acrocnema* in Colorado).

Reconciling Absolute and Relative Population Estimates

For comparing yearly or site-to-site fluctuations in population size, only relative estimates of abundance are needed. Clearly, though, the ability to calibrate these estimates to reflect the underlying absolute abundances is of great benefit. Relative abundance estimates are usually calibrated by comparison with absolute estimates generated under the same conditions (often on the same day). It is during comparison and especially calibration that the assumptions of relative methods are important. The principal ones are:

1. Either the worker's searching efficiency does not vary in time and/or space, or appropriate allowances can be made
2. Either all individuals are equally sightable, or the sightable fraction remains reasonably constant (analog of "equal catchability")
3. Sighting an individual does not alter the probability of sighting another—or the same individual again, if censuses are repeated frequently (analog of "marking has no effect").

Douwes (1970, 1976), Pollard (1977), and Thomas (1983) have demonstrated that counts of individuals concord highly with population size estimates from MRR models, and calibration can therefore be done

by linear regression (Douwes used the Jolly-Seber model for calibration, Pollard and Thomas several frequency-of-capture models). Thomas (1983) examines the question of calibration in greatest detail, demonstrating that the linearity holds over a rather large range of population sizes. All these authors also nicely document species-specific variation in sightability. For example, Douwes (1976) consistently counted about 30 percent of resident *Heodes virgaureae* and *Clossiana selene* on his transects, whereas Pollard (1977) could count nearly all the *Coenonympha pamphilus*, but less than 25 percent of the *Aphantopus hyperantus*.

These studies on calibrating transect counts are highly encouraging (see also the contributions to line-transect theory made by Gates, 1969, and Sen et al., 1974). Thomas (1983, p. 209) has a cogent argument in that "transect recording [may be] a more accurate way of estimating the numbers of a species that flies infrequently or has large populations, for the recapture rate of marked individuals is then so low that traditional methods yield very poor results." (The question is one of sampling intensity; note that the Fisher-Ford model performs well with scanty data). However, reliance on frequency-of-capture models for calibration is questionable. The problem has two aspects.

First, some field studies (Singer and Wedlake, 1981; Gall, 1984b, and unpublished) on the effects of marking butterflies have shown that, contrary to the investigator's intentions and impressions, even careful marking can perturb subsequent activity. When such mark effects do occur, they are virtually always in the direction of depressing flight activity. Marked butterflies are thus at reduced capture risk, at least temporarily, to both unmarked ones and previously marked ones which have had sufficient time to recover. Because frequency-of-capture models rely upon rapid re-mixing of marks back into the resident population, there will be fewer recaptures than expected. A positive bias in $\hat{N}_i$ results, which in some instances can be as much as double or triple the true population size (see Gall, 1984b, for elaboration). Such bias likely will not greatly confound comparison of *trends* in population size among several sites or years, but it *will* prevent accurate calibration. For example, if mark effects introduce bias linearly over a range of population sizes (as seems reasonable for Lepidoptera), then the slopes of calibration equations will not be affected, but the intercepts will be. This also underscores the point that a marking effect's influence is often invisible unless one makes an explicit test for its presence—which is done surprisingly infrequently.

Second, frequency-of-capture methods have several intrinsic shortcomings. Many generalized truncated distributions can be found which fit the observed recapture classes; however, the unobserved zero-classes implied by these distributions vary widely. Because estimating population size by frequency-of-capture involves sums over all recapture classes, different N_i values can be obtained simply by selecting different truncated distributions. Moreover, these truncated distributions are really

descriptions, not models, and so there is usually no strong rationale for choosing one over another. Cormack (1979, p. 231), addressing Efron and Thisted's (1976) MRR study of Shakespearian text, makes the point succinctly: "different models, wholly consistent with the observed data, give totally different estimates of the population size, even when the observations comprise 31,534 individuals observed in total 884,647 times."

I must again stress that criticism of any abundance model is always context dependent, and the above only pinpoints problems inherent in using frequency-of-capture models to calibrate transect counts, *not* problems inherent in transect counts themselves. The distinction is not trivial—lest the reader take home the wrong message—for transect and other direct counts offer perhaps the simplest, most robust, and least expensive (in the broad sense) methods for indexing butterfly abundance.

DISCUSSION

Historical Impact of Lepidopteran Mark-Recapture Research

Lepidoptera have always occupied prominent positions in the theory and practice of most branches of ecology and evolutionary biology. Mark-recapture is no exception to the rule.

The Fisher-Ford, Jackson, and Schnabel methods are the forebears to all subsequent multiple-marking absolute abundance models. Notably, the Fisher-Ford model was conceived, refined, and tested using Lepidoptera as the study organisms. The initial research included population surveys of the lycaenid, *Polyommatus icarus* (Dowdeswell et al., 1940), the arctiid, *Panaxia dominula* (Fisher and Ford, 1947), and the satyrid, *Maniola jurtina* (Dowdeswell et al., 1949). Indeed, Lepidoptera have greatly influenced the development of most of the more prominent MRR models. As with Fisher-Ford, the Craig (1953) and Manly and Parr (1968) models were conceived with Lepidoptera in mind, the former author using data on *Colias*, the latter data on *Zygaena*. Jolly (1982) uses lepidopteran data for the worked examples of his new models, and the standard texts by Southwood (1978), Begon (1979) and Blower et al. (1981) are illustrated with many such lepidopteran examples.

But the frequent use of butterflies and moths for mark-recapture has had more telling impact on science. The mark-recapture research on *Panaxia dominula* by the British ecological geneticists (using the Fisher-Ford model) is of huge historical significance. These hallmark population studies inaugurated heated transcontinental dialogues on the roles of natural selection and genetic drift in natural populations (e.g., Wright, 1948; [pointed] summary by Ford, 1975). This "selectionist-neutralist" debate has occupied a central role in the development of evolutionary theory ever since. It remains a lively subject today, recast with respect to the significance of electrophoretically detectable enzyme variation, and

reevaluation of neo-Darwinian evolutionary tenets (especially pan-selectionism; see Lewontin and Gould, 1979).

Three other long term mark-recapture studies of Lepidoptera deserve mention in this context. First, Kettlewell and his colleagues have used MRR to examine the operative forces underlying the phenomonon of industrial melanism, mostly with noctuid and geometrid moths (summary by Kettlewell, 1973). Second, the 20+ year studies of checkerspot butterflies by Ehrlich et al. have sparked wide debate on the importance of gene flow, and the nature of populations as evolutionary units (historical perspective on *Euphydryas editha* by Ehrlich et al., 1975; Brown and Ehrlich, 1980, for *E. chalcedona*). Third, the studies by Watt's group on Nearctic *Colias* butterflies during the last decade successfully link results from some of the most rigorous mark-recapture work with those from microevolutionary genetics. Their biochemical (Watt, 1983; Watt el al., 1983) and populational (Watt et al., 1977, 1979) articles detailing the action of natural selection on structural gene polymorphisms set the current standard for the discipline.

Thoughts for the Future

What, then, remains to be done? Regarding mark-recapture theory in general, there are still four major needs, among others: (1) models free from the assumption of independence of successive samples, an assumption breached by marking effects in their broadest sense; (2) models to measure local movement patterns, and an interface from these to MRR models; (3) means for dealing with very large populations; and (4) comprehensive models with greater parsimony (Jolly, 1982, has taken the major step in this direction).

Topic two, measuring local movement, touches upon a problem central to all biological field work—defining the limits of populations. Mark-recapture is the technique for measuring movement patterns, and I expect that studies of Lepidoptera will contribute significantly to advances in this area, as they have in the development of MRR models (already, butterflies have figured in recent innovative work on local movement: e.g., Jones et al., 1980; Kareiva, 1982). Moreover, such mark-recapture studies will prove of great value to conservationists, for local movement patterns define (1) the physical boundaries of a population, (2) its interconnectedness with other populations, and hence (3) effective neighborhood sizes. Mark-recapture also provides data on (4) effective population sizes, and in combination these four factors are relevant to understanding extinction and colonization probabilities, the general subject of which is reviewed by Soule and Wilcox (1980). Similarly, the first two factors are useful in the design and maintenance of preserves. For example, the movements of individuals pinpoint the location of dispersal corridors, and hence offer insight into how to partition parcels slated for

development i.e., don't build across the major dispersal corridors (opportunity for re-colonization would otherwise be drastically reduced).

As aptly noted by Murphy (1984), in a tart review of Arnold's (1983) mark-recapture surveys of endangered California butterflies, our Palearctic counterparts have amassed embarassingly large leads in the practice and politics of invertebrate conservation. This is quite true, and it is novel mark-recapture research that will go far toward establishing parity. Thus, although some of Murphy's caveats deserve to be heeded, I cannot recommend in the least his assertion (p. 268) that: "in essence the results of mark-recapture studies, no matter how rigorous, and natural history investigations, no matter how detailed, by themselves tell us virtually nothing at all about the extinction vulnerability in butterflies."

Age structure in adult Lepidoptera is a subject well deserving of final mention. The usual index of butterfly age is physical wing wear—most often scale loss, but sometimes cuticular damage. While correlations between age and indices of one or both aspects of wear have been independently discovered and reported many times, only recently has age structure been integrated quantitatively into lepidopteran MRR studies. Watt et al. (1977, 1979) and Tabashnik (1980) were the first to do so, focusing principally on age-specific patterns in residence. Extension of this research into catchability and movement has been productive, with perhaps the most intriguing finding being that sex-age-specific movement may be a widespread pattern, highlighted by emigration of old females (see Gall, 1984a, 1984b, 1984c).

A sobering conclusion from these MRR studies is that failure to consider age structure often entirely masks crucial biological patterns. The extent to which age, when not treated, confounds the results of most published studies of lepidopteran biology is as yet unclear, but I judge it to be a potentially explosive problem (Gall, 1984c, for discussion). Age-specific movement also bears directly on the needed refinements in MRR theory outlined above, and so I feel that a quantitative framework for dealing with lepidopteran age structure is likely to be among the more important future contributions to the field of mark-recapture. Because lepidopteran age can be indexed so simply, I am also optimistic for rapid progress in this area.

Literature Cited

ARNOLD, R. A., 1983. Ecological studies of six endangered butterflies (Lepidoptera, Lycaenidae): island biogeography, patch dynamics, and the design of habitat preserves. Univ. Calif. Publ. Entom. 99:1-161.

BAILEY, N. T. J., 1951. On estimating the size of mobile populations from capture-recapture data. Biometrika 38:293-306.

————, 1952. Improvements in the interpretation of recapture data. J. Anim. Ecol. 21:120-127.

BEGON, M., 1979. Investigating Animal Abundance: Capture-Recapture for Biolo-

gists. University Park Press, Baltimore. 97 pp.

BISHOP, J. A. & P. M. SHEPPARD, 1973. An evaluation of two capture-recapture models using the technique of computer simulation. Pages 235-252, *In*: M. S. Bartlett and R. W. Hiorns (eds.), The Mathematical Theory of the Dynamics of Biological Populations. Academic Press, London.

BLOWER, J. G., L. M. COOK & J. A. BISHOP, 1981. Estimating the Size of Animal Populations. Alden Press, London. 128 pp.

BROWN, I. L. & P. R. EHRLICH, 1980. Population biology of the checkerspot butterfly, *Euphydryas chalcedona*: structure of the Jasper Ridge colony. Oecologia 47:239-251.

BRUSSARD, P. F., 1970. Field techniques for investigations of population structure in a "ubiquitous" butterfly. J. Lepid. Soc. 25:22-29.

CAROTHERS, A. D., 1973. Capture-recapture methods applied to a population with known parameters. J. Anim. Ecol. 42:125-146.

CAUGHLEY, G., 1977. Analysis of Vertebrate Populations. Wiley, London. 234 pp.

CORMACK, R. M., 1968. The statistics of capture-recapture. Ann Rev. Oceanogr. Marine Biol. 6:455-506.

_____, 1979. Models for capture-recapture. Pages 217-255, *In*: R. M. Cormack, G. P. Patil, and D. S. Robson (eds.), Sampling Biological Populations. International Co-operative Publishing House, Fairview, Maryland.

CRAIG, C. C., 1953. On the utilisation of marked specimens in estimating populations of flying insects. Biometrika 40:170-176.

DOANE, C. L. & M. L. McMANUS, 1981. The gypsy moth: research toward integrated pest management. USDA Tech. Bull. 1584:1-757.

DOUWES, P., 1970. Size of, gain to and loss from a population of *Heodes virgaureae* L. (Lep., Lycaenidae). Entom. Scand. 1:263-281.

_____, 1976. An area census method for estimating butterfly population numbers. J. Res. Lepid. 15:146-152.

DOWDESWELL, W. H., R. A. FISHER & E. B. FORD, 1940. The quantitative study of populations in the Lepidoptera. 1. *Polyommatus icarus*, Rott. Ann. Eugenics 10: 123-136.

_____, 1949. The quantitative study of populations in the Lepidoptera. 2. *Maniola jurtina*, L. Heredity 3:67-84.

EDWARDS, W. R. & L. L. EBERHARDT, 1967. Estimating cottontail abundance from live-trapping data. J. Wildl. Mgmt. 31:87-96.

EFRON, B. & R. THISTED, 1976. Estimating the number of unseen species: how many words did Shakespeare know? Biometrika 63:435-447.

EHRLICH, P. R. & S. E. DAVIDSON, 1960. Techniques for capture-recapture of Lepidoptera populations. J. Lepid. Soc. 14:227-229.

_____, R. R. WHITE, M. C. SINGER, S. W. McKECHNIE & L. E. GILBERT, 1975. Checkerspot butterflies: an historical perspective. Science 188:221-228.

FISHER, R. A. & E. B. FORD, 1947. The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula* (L.). Heredity 1:143-174.

FORD, E. B., 1975. Ecological Genetics. Methuen, London. 442 pp.

GALL, L. F., 1984a. Population structure and recommendations for conservation of the narrowly endemic alpine butterfly, *Boloria acrocnema* (Lepidoptera: Nymphalidae). Biol. Conservation: 28:111-138.

_____, 1984b. The effects of capturing and marking on subsequent activity in *Boloria acrocnema* (Lepidoptera: Nymphalidae), with a comparison of dif-

ferent numerical models that estimate population size. Biol. Conservation: 28:139-154.

————, 1984c. The evolutionary ecology of a species-rich sympatric array of *Catocala* moths. Ph.D. Dissertation, Yale University. 264 pp.

GATES, C. E., 1969. Simulation study of estimators for the line transect sampling method. Biometrics 25:317-328.

HUGHES, S. M., 1975. Fourth of July butterfly count. Xerces Society Self-Help Sheet #3 (15 May 1975). 4 pp.

JOLLY, G. M., 1965. Explicit estimates from capture-recapture data with both death and immigration - stochastic model. Biometrika 52:225-247.

————, 1982. Mark-recapture models with parameters constant in time. Biometrics 38:301-321.

JONES, R. E., N. GILBERT, M. GUPPY & V. NEALIS, 1980. Long-distance movement of *Pieris rapae*. J. Anim. Ecol. 49:629-642.

KAREIVA, P., 1982. Experimental and mathematical analyses of herbivore movement: quantifying the influence of plant spacing and quality on foraging discrimination. Ecol. Monogr. 52:261-282.

KELKER, G. H., 1940. Estimating deer populations by a differential hunting loss in the sexes. Proc. Utah Acad. Sci., Arts. Lett. 17:65-69.

KETTLEWELL, H. B. D., 1973. The Evolution of Melanism. Clarendon Press, Oxford. 423 pp.

GOULD, S. J. & R. C. LEWONTIN, 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc. Roy. Soc. Lond. B 205:581-598.

LeCREN, E. D., 1965. A note on the history of mark-recapture population estimates. J. Anim. Ecol. 34:453-454.

MANLY, B. F. J., 1970. A simulation study of animal population estimation using the capture-recapture method. J. Appl. Ecol. 7:13-39.

———— & M. J. PARR, 1968. A new method of estimating population size, survivorship, and birth rate from capture-recapture data. Trans. Brit. Soc. Entom. 18:81-89.

MURPHY, D. D., 1984. Review of: Ecological studies of six endangered butterflies (Lepidoptera, Lycaenidae): island biogeography, patch dynamics, and the design of habitat preserves (by Arnold, R. A., 1984). J. Res. Lepid. 22:267-269.

———— & P. R. EHRLICH, 1980. Two California checkerspot butterfly subspecies: one new, one on the verge of extinction. J. Lepid. Soc. 34:316-320.

OPLER, P. A. & J. A. POWELL, 1984. Butterfly Counts, 1982 and 1983 summaries. The Xerces Society. 43 pp.

POLLARD, E., 1977. A method for assessing changes in the abundance of butterflies. Biol. Conservation 12:115-134.

————, 1979. A national scheme for monitoring the abundance of butterflies: the first three years. Brit. Entom. Nat. Hist. Soc., Proc. and Trans. 12:77-90.

————, 1984. Fluctuations in the abundance of butterflies. Ecol. Entomol. 9: 179-188.

————, D. O. ELIAS, M. J. SKELTON & J. A. THOMAS, 1973. A method of assessing the abundance of butterflies in Monks Wood National Nature Reserve in 1973. Entom. Gaz. 26:76-88.

PYLE, R. M., 1976. Conservation of Lepidoptera in the United States. Biol. Conservation 9:55-75.

————, M. BENTZIEN & P. A. OPLER, 1981. Insect conservation. Ann. Rev. Entom. 26:233-258.

ROFF, D. A., 1973. On the accuracy of some mark-recapture estimators. Oecologia 12:35-54.

SEBER, G. A. F., 1965. A note on the multiple-recapture census. Biometrika 52:249-259.

————, 1973. The Estimation of Animal Abundance and Related Parameters. Griffin Press, London. 506 pp.

SEN, A. R., J. TOURIGNY & G. E. J. SMITH, 1974. On the line transect sampling method. Biometrics 30:329-340.

SINGER, M. C. & P. WEDLAKE, 1981. Capture does affect probability of recapture in a butterfly species. Ecol. Entom. 6:215-216.

SMITH, L. V., 1984. A twelve year count of three California butterflies. J. Lepid. Soc. 37:275-280.

SOULE, M. E. & B. E. WILCOX, 1980. Conservation Biology. Sinauer Assoc., Sunderland. 395 pp.

SOUTHWOOD, T. R. E., 1978. Ecological Methods (2nd ed.). Halsted Press, New York. 524 pp.

TABASHNIK, B. E., 1980. Population structure of pierid butterflies. III. Pest populations of *Colias philodice eriphyle.* Oecologia 47:175-183.

THOMAS, J. A., 1983. A quick method for estimating butterfly numbers during surveys. Biol. Conservation 27:195-211.

WATT, W. B., 1983. Adaptation at specific loci. II. Demographic and biochemical elements in the maintenance of the *Colias* PGI polymorphism. Genetics 103: 691-724.

————, R. C. CASSIN & M. S. SWAN, 1983. Adaptation at specific loci. III. Field behavior and survivorship differences among *Colias* PGI genotypes are predictable from *in vitro* biochemistry. Genetics 103:725-739.

————, F. S. CHEW, L. G. SNYDER, A. G. WATT & D. E. ROTHSCHILD, 1977. Population structure of pierid butterflies. I. Numbers and movements of some montane *Colias* species. Oecologia 27:1-22.

————, D. HAN & B. E. TABASHNIK, 1979. Population structure of pierid butterflies. II. A "native" population of *Colias philodice eriphyle* in Colorado. Oecologia 44:44-52.

WRIGHT, S. A., 1948. On the roles of directed and random changes in gene frequency in the genetics of populations. Evolution 2:279-294.